Genetica (2009) 136:141–147
DOI 10.1007/s10709-008-9327-x

Assessing genetic diversity in *Gossypium arboreum* L. cultivars using genomic and EST-derived microsatellites

Stella K. Kantartzi · Mauricio Ulloa · Erik Sacks · James McD. Stewart

Received: 11 October 2007 / Accepted: 25 September 2008 / Published online: 14 October 2008

Abstract The cultivated diploid, *Gossypium arboreum* L., (A genome) is an invaluable genetic resource for improving modern tetraploid cotton (*G. hirsutum* L. and *G. barbadense* L.) cultivars. The objective of this research is to select a set of informative and robust microsatellites for studying genetic relationships among accessions of geographically diverse *G. arboreum* cultivars. From more than 1,500 previously developed simple sequence repeat (SSR) markers, 115 genomic (BNL) and EST-derived (MUCS and MUSS) markers were used to evaluate the allelic diversity of a core panel of *G. arboreum* accessions. These SSR data enabled advanced genome analyses. A set of 25 SSRs were selected based both upon their high level of informativeness (PIC $\geq$ 0.50) and the production of clear PCR bands on agarose gels. Subsequently, 96 accessions representing a wide spectrum of diversity of *G. arboreum* cultivars were analyzed with these markers. The 25 SSR loci revealed 75 allelic variants (polymorphisms) ranging from 2 to 4 alleles per locus. The Neighborjoining (NJ) method, based on genetic dissimilarities, revealed that cultivars from geographically adjacent countries tend to cluster together. Outcomes of this research should be useful in decreasing redundancy of effort and in constructing a core collection of *G. arboreum*, important for efficient use of this genetic resource in cotton breeding.

Keywords Cotton · Diversity · EST-SSRs · Genomic-SSR · *Gossypium arboreum*

S. K. Kantartzi (✉)
Southern Illinois University, Plant Soil and Agricultural Systems, Carbondale, IL 62901, USA
e-mail: kantart@siu.edu

M. Ulloa
USDA-ARS, Western Integrated Cropping Systems, Shafter, CA 93263, USA

E. Sacks
USDA-ARS, Stoneville, MS 38776, USA

J. McD. Stewart
Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR 72701, USA

Introduction

Cotton (*Gossypium* L.) is the leading fiber crop in the world. Although the genus has approximately 50 species, including 45 diploids (A through G plus K) and five allotetraploids (AD)(Fryxell 1992; Fryxell et al. 1992), only four of them are cultivated, including two diploid ($2n = 2\times = 26$): *G. arboreum* L. (A_2A_2) and *G. herbaceum* L. (A_1A_1), and two tetraploid ($2n = 4\times = 52$): *G. hirsutum* L. (AD_1AD_1) and *G. barbadense* L. (AD_2AD_2) species. The narrow genetic base of the primary cotton (*G. hirsutum*) breeding gene pool is one of the major constraints in breeding programs worldwide. This underscores the necessity of enriching the gene pool with genetic diversity (Abdalla et al. 2001). The study of genetic relationships between strains of diploid cotton from various ecological regions, provides not only a theoretical basis for conserving diploid cotton germplasm resources, but also for targeting and improving certain ideal characteristics such as early maturity, resistance to stress, and fiber quality, and for exploiting this secondary germplasm pool in modern cotton production (Stewart 1995; Guo et al. 2003; Ulloa et al. 2007).

Genetic diversity is desirable for long-term crop improvement and reduction of vulnerability to important crop stresses. Many successful cotton cultivars have been

developed from closely related parents, but limited yield gains in recent years have led some to advocate more extensive use of exotic germplasm (Meredith 1991). Unless methods are improved to transfer useful allelic variation from diverse germplasm resources to the primary cotton breeding gene pools, cotton germplasm resources will remain largely underutilized (Van Esbroeck et al. 1997). Stewart (1995) presented several strategies for efficient utilization of the germplasm pools of cotton. Transfer of desirable genes through introgression from germplasm resources of other *Gossypium* species could play an important role in decreasing genetic variability in cultivated *G. hirsutum*. And, molecular markers are expected to increase the efficiency with which traits can be introgressed (Ulloa et al. 2007).

Gossypium arboreum possesses many favorable traits for cotton production which the tetraploid cultivars lack. For example, drought tolerance, resistance to diseases like black root rot, and pests such as reniform nematodes and aphids, make *G. arboreum* well adapted to dry land conditions and low input cultivation practices. Natural *G. arboreum* fibers display various colors (e.g., white, milky, beige and brown). Some of the accessions produce fibers with high strength and seeds with high oil content and seed index (Mehetre et al. 2003). The *G. arboreum* germplasm collection is an invaluable genetic resource for tetraploid cotton improvement; however, it has not been well characterized at the molecular level. Understanding the genetic relationships within *G. arboreum* would facilitate efficient use of this resource for developing superior cotton cultivars with favorable agronomic traits.

Molecular markers have been extensively used to study genetic diversity, genetic relationships and molecular phylogeny in *Gossypium* species. Especially useful have been the PCR-based markers such as random amplified polymorphic DNA (RAPD) (Tatineni et al. 1996; Xu et al. 2001; Lu and Myers 2002), amplified fragment length polymorphism (AFLP) (Abdalla et al. 2001; Alvarez and Wendel 2006), and simple sequence repeats aka microsatellites (SSRs) (Lui et al. 2000b; Zhu et al. 2003).

Within recent years, SSRs have become one of the most powerful genetic marker systems in biology with great potential for application in plant breeding programs (Gupta and Varshney 2000). They are among the most abundant, easy to use, and variable DNA sequences, i.e, they are highly polymorphic. For these reasons SSRs have become an important marker system in cultivar fingerprinting, diversity studies, molecular mapping and in marker-assisted selection (Reddy et al. 2001). The loci of these markers are highly transferable across species (>50%) especially within a genus (Saha et al. 2004). In addition, SSR markers derived from expressed sequence tags (EST-SSRs) are likely to be even more transferable than genomic SSRs because they are a part of the transcribed regions of DNA (Park et al. 2005). Transcribed regions are more conserved across species and genera, thus EST-SSRs can be used for comparative mapping (Saha et al. 2004).

We evaluated EST- and genomic-SSRs for their informativeness by assessing their allelic polymorphism within a collection of *G. arboreum* accessions. The results are expected (1) to allow selection of informative and robust microsatellites for assembly of a 'genotyping set' of SSR primers and (2) to clarify genetic relationships among accessions of geographically diverse *G. arboreum* cultivars.

Materials and methods

The 96 accessions of *G. arboreum* examined for this project, along with their PI numbers and passport data are given in Table 1. A core set of 15 accessions (Table 2) was selected to screen 115 microsatellites for polymorphisms Fresh, young leaves from each *G. arboreum* accession were used to isolate DNA. Approximately 200 mg of tissue per sample were ground in liquid nitrogen and the resulting powder mixed in a microtube with the extracting buffer supplied with a DNeasy® Plant Mini Kit (Qiagen Inc., Valencia, CA). All procedures for DNA isolation were according to the manufacturer's protocol for the kit.

The microsatellite primer pairs originated from the following sources: 50 were developed at the Brookhaven National Lab (BNL) (http://www.algodon.tamu.edu/cgibin/ace/searches/browser) from *G. hirsutum* genomic DNA. EST-SSRs primer pairs were developed by Park et al. (2005) from *G. arboreum* cv. AKA 8401of which 65 were selected for further examination (45 for MUSS and 20 for MUCS: M for microsatellite, U last name of principal investigator, SS for simple sequences, and CS for complex sequence repeats). Ultimately 25 primer pairs (Table 3) were selected as being sufficient for determining the molecular diversity of *G. arboreum* accessions (13 MUSS, 11BNL, and 1 MUCS). The details of the microsatellite markers and primer sequences may be found in the Cotton Microsatellite Database at http://www.cottonmarker.org.

PCR amplification was performed in a volume of 50 µl containing 1 µl of DNA extract (50 ng/µl), 250 nM of each primer, 200 µM of dNTPs, 0.5 U of Taq polymerase and 10 µl of reaction buffer. Thirty PCR cycles, each consisting of 30 s denaturation at 96°C, 30 s annealing at [Tm −5°C] (optimum annealing temperature for each primer pair) and 1-min-polymerization at 72°C, were performed in a Hybaid thermocycler (Thermo-Fisher Scientific, Waltham, MA). Polymorphism at each locus was assessed by electrophoresis of the PCR products in a horizontal gel system at 110 V for 4 h through 4% Metaphore gels stained with ethidium

Table 1 Ninety-six *G. arboreum* accessions used in this study with their respective country of origin

Entry	Leaf Sample ID	PI	Origin	Name	Entry	Leaf Sample ID	PI	Origin	Name
1	A2-0018	89202	Pakistan	Sanguinium	49	A2-0159	529776	China	Jia Zihua
2	A2-0019	129723	India	Verum	50	A2-0160	529777	China	Fuyang Zisa
3	A2-0020	129742	India	Malvi 9	51	A2-0161	529778	China	Chengfeng He
4	A2-0021	152088	Uzbekistan	Nanking	52	A2-0162	529779	China	Baoshan Tumi
5	A2-0024	179559	India	Desi Rui	53	A2-0164	529781	China	Shih
6	A2-0027	180243	India	Kapas	54	A2-0169	529786	India	Bani
7	A2-0034	183160	India	Neglectum	55	A2-0174	529791	Uzbekistan	Line 10824 II
8	A2-0042	185786	Korea	Hamju	56	A2-0176	529793	Uzbekistan	Line 0 1423 2
9	A2-0043	185787	Korea	Samju	57	A2-0178	529795	India	Manihur
10	A2-0044	185788	Korea	Yong-Gang	58	A2-0179	529796	Unknown	Wightianum
11	A2-0045	200421	Pakistan	Mollisani	59	A2-0181	529798	India	Verum 434
12	A2-0047	213373	Sudan	Gaorani	60	A2-0186	529803	Rusian Federation	Race Sinense
13	A2-0050	324493	India	Cernuum	61	A2-0187	529804	Rusian Federation	CV 2557
14	A2-0054	408751	Pakistan	Roseum	62	A2-0188	529805	Rusian Federation	Soudanense
15	A2-0057	408754	India	Company No. 3	63	A2-0190	615699	Myanmar	Burma C19
16	A2-0062	408759	Iran	Koresu	64	A2-0191	615700	China	Chinese Narrow Leaf
17	A2-0066	408763	India	Small Sport	65	A2-0192	615701	Iran	Soudanense
18	A2-0067	408764	India	Kokato Narrow	66	A2-0194	615703	India	CJ 73
19	A2-0068	408765	India	Sanguienum 1	67	A2-0195	615704	India	Gaorani 46
20	A2-0069	408766	India	Aligarh A19	68	A2-0229	615730	China	Zi Se Da Hua
21	A2-0072	408769	India	Chin-Wa	69	A2-0230	615731	China	Hong Zie Mian
22	A2-0073	408770	China	Chinese Naked	70	A2-0231	615732	China	Tu Mian
23	A2-0074	408771	China	Chinese Pale	71	A2-0232	615733	China	Zi Hua Guang Zi
24	A2-0078	417886	India	Virnar	72	A2-0234	615734	China	Bai Wan Mian
25	A2-0080	417888	India	Gaorani 46	73	A2-0235	615736	China	Le Ting Xian Hezi
26	A2-0082	417890	India	Gareo	74	A2-0236	615737	China	Guang Di Xia Tie Zi
27	A2-0083	417891	India	Deshi 9	75	A2-0237	615738	China	Zhejiang Ding Hai Zi
28	A2-0084	417892	India	Sanguienum Minor	76	A2-0240	615741	China	Lianhua Zhing Mian
29	A2-0085	417893	India	Chang Fend	77	A2-0241	615742	China	Ging Jing
30	A2-0087	417895	India	Northern 14	78	A2-0245	615746	Pakistan	Red Flower
31	A2-0088	417896	India	Bani 31	79	A2-0249	615750	Pakistan	Punjabi 39
32	A2-0089	417897	India	Coconada 1 20	80	A2-0253	615754	India	Jarilla
33	A2-0090	417998	India	(Red Arboreum) Poona	81	A2-0256	615757	Thailand	Unverified
34	A2-0091	417899	India	Nanking Khakhi	82	A2-0259	408797	India	Anbu 13
35	A2-0100	529728	Unknown	Nanking	83	A2-0260	408798	China	Anc 1
36	A2-0101	529729	Unknown	Garo Hill	84	A2-0263	451745	China	Shantung xiao kwang hsu
37	A2-0102	417911	India	Dudhatal Yellow	85	A2-0264	451746	China	Ting shen xiao pai kai mien
38	A2-0111	529738	Unknown	Okinawa	86	A2-0265	529644	India	Wagad 14
39	A2-0141	452093	China	Chang Zi No.1	87	A2-0300	615794	India	Indicum 2
40	A2-0142	452094	China	Liaoyang No.1	88	A2-0301	615795	India	Indicum 10
41	A2-0143	452095	China	Millionaire	89	A2-0986	616480	India	White Kapas
42	A2-0144	452096	China	Nantong	90	A2-1005	629333	India	Gao CP 3
43	A2-0145	452097	China	Shanghai Ying	91	A2-1011	629339	India	Kanpur

Table 1 continued

Entry	Leaf Sample ID	PI	Origin	Name	Entry	Leaf Sample ID	PI	Origin	Name
44	A2-0146	452098	China	Zejiang Yuyao	92	A2-1096	629424	India	Arvensis
45	A2-0153	417885	India	Baluhistan 7	93	A2-1097	629425	India	Adonicum
46	A2-0154	441979	China	Liaochung 6	94	A2-1130	629458	India	Burma Silk
47	A2-0156	529773	China	Jiang Yin	95	A2-1244	629572	India	Deshi 8
48	A2-0158	529775	China	Shixiya 1	96	A2-1596	629924	India	Obtusifolium Hirsutu

Table 2 Core set of *G. arboreum* plants that were used for the screening 115 microsatellite primers

Entry	Leaf Sample ID	PI	Origin	Name
1	A2-0019	129723	India	Verum
2	A2-0021	152088	Uzbekistan	Nanking
3	A2-0027	180243	India	Kapas
4	A2-0045	200421	Pakistan	Mollisani
5	A2-0043	185787	Korea	Samju
6	A2-0100	529728	Unknown	Nanking
7	A2-0186	529803	Rusian Federation	Race Sinense
8	A2-0187	529804	Rusian Federation	CV 2557
9	A2-0190	615699	Myanmar	Burma C19
10	A2-0192	615701	Iran	Soudanense
11	A2-0234	615734	China	Bai Wan Mian
12	A2-0240	615741	China	Lianhua Zhing Mian
13	A2-0253	615754	India	Jarilla
14	A2-0256	615757	Thailand	Unverified
15	A2-0264	451746	China	Ting shen xiao pai kai mien

bromide (0.5 μg ml^{-1}) using 1xTBE running buffer (89 mM Tris, 89 mM borate, 2 mM EDTA pH 8.3). Finally, the gels were photographed under UV light attached to a gel documentation system (Bio-Rad, Hercules, CA).

The profiles produced by SSR markers were scored manually: each allele was scored as present (1) or absent (0) for each of the SSR loci. Genetic diversity was calculated at each locus for allelic polymorphism information content (PIC) (Anderson et al. 1993), with program CERVUS version 2.0 based on allelic frequencies among all 96 genotypes analyzed. PIC values of each locus were calculated as: PIC$j = 1 - \sum p^2 lj$, plj is the frequency if the lth allele for locus j and is summed over its L alleles. Markers were classified as informative when PIC $\geq$ 0.5. The data matrix was also converted to a matrix of similarity values (F) using the formula: $F = N_{xy}/(N_x + N_y)$ where N_x and N_y are the numbers of fragments in genotypes x and y, respectively, and N_{xy} is the number of fragments shared by genotypes x and y (Nei and Li 1979). Dissimilarity values ($D = 1 - F$), were used in further calculations. A dendrogram was constructed using the neighbor joining (NJ) method (Saitou and Nei 1987). All computations were carried out with appropriate procedures of the software package NTSys 2.1 (Rohlf 2000).

Results and discussion

One hundred and fifteen SSR primer pairs were used to amplify genomic DNA fragments from the core set of plants. Twenty of the 115 SSR primer sets (15 BNLs, 4 MUCS and 1 MUSS) were not exploited because of poor quality of amplification, a result that was not surprising since all of the BNL SSR primers were developed from the tetraploid cotton genome, while the polymorphisms found in the MUSS and MUCS series involved the two cultivated tetraploids species and two diploid D-genome species as well as the two A-genome species. Liu et al. (2000a) estimated that annealing between primers and templates would be affected when tetraploid cotton-derived primers were used to amplify templates from diploid A or D genome species. Difficult amplification was observed for four MUCS and one MUSS, as a result they were discarded from the study. The other primer sets yielded a total of 95

Table 3 A set of 25 SSRs, sufficient for studies of genetic diversity of *G. arboreum* accessions

#	Primer ID	Repeat Code	Chromosome	No. of alleles	PIC
1	BNL 0448	$(CT)_{13}$	20	4	0.55
2	BNL 0530	$(GA)_{10}$	4	3	0.53
3	BNL 0598	$(CT)_{12}$	12	3	0.66
4	BNL 0632	$(AG)_{13}$	Unmapped	3	0.66
5	BNL 0834	$(CA)_{13}$	17	4	0.59
6	BNL0852	$(CA)_{13}$	5/D08	2	0.50
7	BNL0946	$(GA)_{14}$	20	3	0.64
8	BNL 1162	$(GA)_{14}$	9	3	0.54
9	BNL 1395	$(AT)_{11} + (AG)_{10}$	16	3	0.54
10	BNL 1408	$(AG)_{17}$	11/A03/D02	3	0.56
11	BNL1421	$(AG)_{29}$	18/A01	2	0.50
12	MUCS 0212	$(TCAC)_2$ $(TTTCC)_2$	Unmapped	3	0.57
13	MUSS0011	$(AT)_{27}$	Unmapped	2	0.50
14	MUSS0020	$(AT)_{13}$	Unmapped	3	0.52
15	MUSS0026	$(AT)_{17}$	12	3	0.60
16	MUSS0027	$(TGA)_6$	Unmapped	3	0.51
17	MUSS0049	$(ATA)_{20}$	Unmapped	3	0.65
18	MUSS0083	$(TTTTC)_4$	9	3	0.67
19	MUSS0096	$(TA)_{19}$	10/20	3	0.58
20	MUSS0111	$(TTA)_{15}$	Unmapped	2	0.50
21	MUSS0121	$(CT)_{13}$	Unmapped	3	0.56
22	MUSS0257	$(TA)_{10}$	Unmapped	3	0.60
23	MUSS0300	$(ATC)_4$	23	4	0.74
24	MUSS0321	$(TATTT)_3$	Unmapped	4	0.75
25	MUSS0387	$(GA)_6$	Unmapped	3	0.55

loci and 233 alleles. However, of the 233 alleles, only 218 were polymorphic and were amplified by 80 SSR primers. The mean number of alleles per locus was 2.40 (StDev 0.94), but the number varied from 2 to 5. The PIC values ranged from 0.00 to 0.68 (average 0.42). In Liu et al. (2000b), the PIC, calculated by the same formula, ranged from 0.05 to 0.82 (average 0.31), while the values obtained by Lacape et al. (2007) ranged from 0.08 to 0.89 (average 0.55). The relation between PIC and the number of alleles is shown in Fig. 1.

We propose a set of SSR markers (Table 3) for use in A_2 species diversity studies. The loci were selected based both upon their high level of informativeness (PIC $\geq$0.50) and the production of distinct bands in the metaphor gels. The set of 25 SSRs is sufficient for studies of genetic diversity of *G. arboreum* accessions. Also, identifying a genotyping set of highly polymorphic microsatellite markers is expected to increase the efficiency of genetic diversity assessment and variety identification (Song et al. 1999; Macaulay et al. 2001; Masi et al. 2003; Jain et al. 2004).

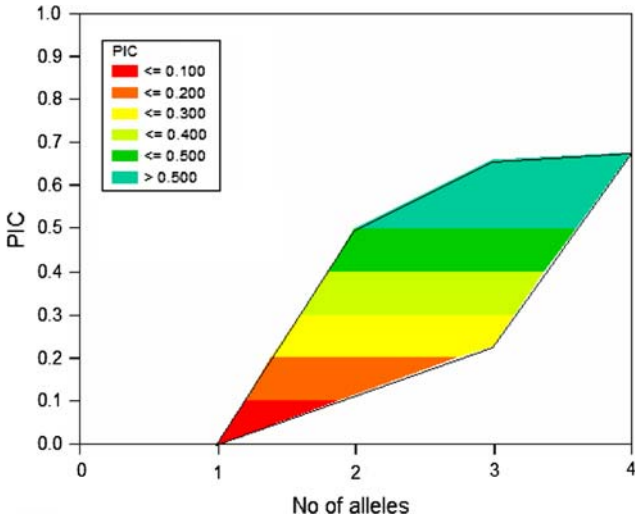

Fig. 1 Plot of PIC statistics as a function of the number of alleles for the SSRs analyzed across the core set of *G. arboreum* accessions

Eleven BNLs (22%), one MUSC (5%) and thirteen MUSS (29%) (Table 3) had PIC values of 0.50 or greater. Within these 25 informative loci, 72% of the SSRs had a di-nucleotide motif, 16% had a tri-nucleotide motif and the remaining 12% had tetra/penta-nucleotide repeats, indicating that the higher frequency of informative SSRs for *G. arboreum* had di-nucleotide motifs. In our study, the majority (80%) of the informative SSRs contained at least 10 repeats. Although contradictory references also exist (e.g., Struss and Plieske 1998), a similar positive relationship between the number of tandem repeats and the level of polymorphism also was observed in tomato (Smulders et al. 1997) and maize (Vigouroux et al. 2002).

The NJ tree (Fig. 2) demonstrates the relationships between the 96 *G. arboreum* cultivars examined. The accessions from China, E and SE Asian countries (Thailand, Korea and Myanmar) and the ones from India, S and Central Asia (Pakistan, Uzbekistan, Iran) were clustered in two divergent groups. Similar results were found by Vergara and Bughrara (2003) who reported that bentgrass accessions from geographically adjacent countries commonly clustered together. Also, it should be noted that the countries of the former USSR and Iran probably "borrowed" genetic material both from India and China.

Genetic distance coefficients among all 96 *G. arboreum* accessions ranged from 0.005 to 0.510. The minimum (0.005) was between two accessions from Pakistan (A_2-0245 and A_2-0249), which means they were closely related. The maximum coefficient of genetic distance (0.51) was generated between A_2-0073 and A_2-0090. The former was introduced from India and the latter from China. The one from India was outstanding because its average distance to other accessions in the same group was 0.18, showing large

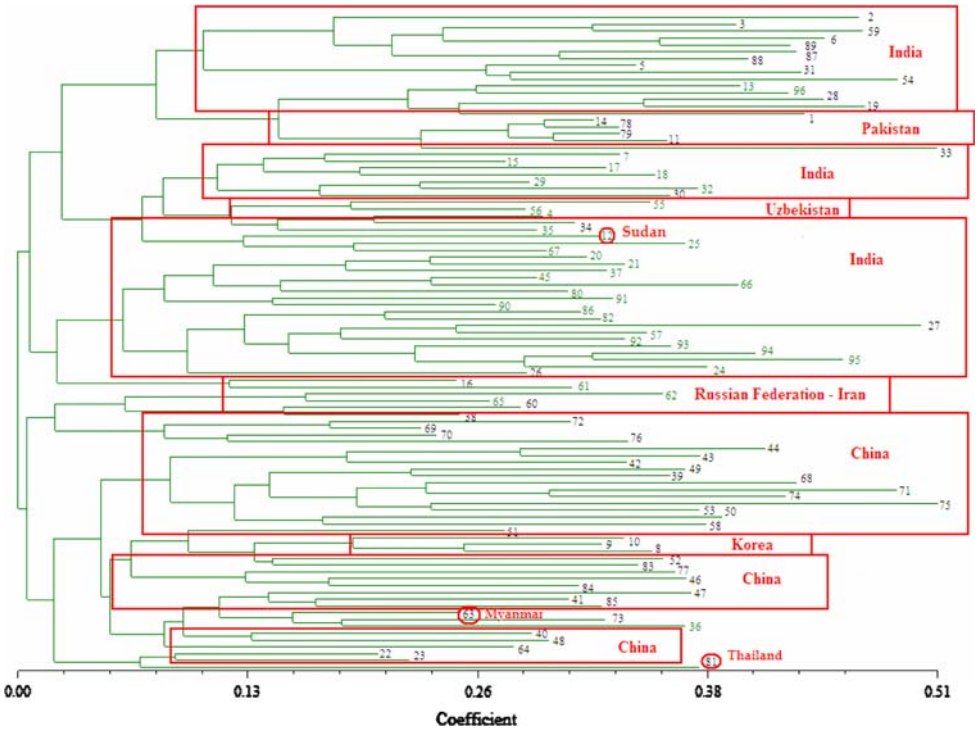

Fig. 2 A phylogenetic NJ tree generated by NTSYSpc2.1 software. Numbers represent the 96 *G. arboreum* accessions described in Table 1

variability in genomic constitution. A group of Chinese cultivars was the most divergent, with a maximum distance index of 0.278 between accessions A_2-230 and A_2-237. *Gossypium arboreum* has been cultivated in China for over 2000 years so it is likely that the widely different environments of China contributed significantly to the diversity of *G. arboreum* (Liu et al. 2006).

A narrow genetic base in cotton (*G. hirsutum*) breeding germplasm (Multani and Lyon 1995; Iqbal et al. 1997; Ulloa et al. 1999; Bowman 2000) and *Gossypium* species (Iqbal et al. 2001; Abdalla et al. 2001) has been reported. *Gossypium arboreum* is an invaluable genetic resource for improving modern cotton cultivars. Molecular markers that reveal DNA polymorphisms have proven to be a powerful tool for genotyping and estimating genetic diversity. Among molecular markers, SSR markers efficiently and effectively reveal the genetic diversity of crop germplasm collections (Khan et al. 2000; Guo et al. 2003). A systematic genetic assessment of the gene resources will help to decrease redundancy and to construct a core germplasm collection, which is important for efficient use of these genetic resources in cotton breeding.

Acknowledgement We thank the Delta & Pine Land Company for financial support.

References

Abdalla AM, Reddy OUK, El-Zik KM, Pepper AE (2001) Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor Appl Genet 102:222–229. doi:10.1007/s001220051639

Alvarez I, Wendel JF (2006) Cryptic interspecific introgression and genetic differentiation within *Gossypium aridum* (Malvaceae) and its relatives. Evol Int J Org Evol 60:505–517

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186. doi:10.1139/g93-024

Bowman DT (2000) Attributes of public and private cotton breeding programs. J Cotton Sci 4:130–136

Fryxell PA (1992) A revised taxonomic interpretation of *Gossypium* L. (Malvaceae). Rheedea 2:108–165

Fryxell PA, Craven LA, Stewart JMD (1992) A revision of *Gossypium* sect. *Grandicalyx* (Malvaceae), including the description of six new species. Syst Bot 17:91–114. doi:10.2307/2419068

Guo WZ, Wang K, Zhang TZ (2003) A and D genome evolution in *Gossypium* revealed using SSR molecular markers. Acta Genet Sin 30:183–188

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185. doi:10.1023/A:1003910819967

Iqbal MJ, Aziz N, Saeed NA, Zafar Y, Malik KA (1997) Genetic diversity evaluation of some elite cotton varieties by RAPD-analysis. Theor Appl Genet 94:139–144. doi:10.1007/s001220050392

Iqbal MJ, Reddy OUK, El-Zik KM, Pepper AE (2001) A genetic bottleneck in the evolution under domestication of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. Theor Appl Genet 103:547–554. doi:10.1007/PL00002908

Jain S, Jain RK, McCouch S (2004) Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently-labeled microsatellite markers. Theor Appl Genet 109:965–977. doi:10.1007/s00122-004-1700-2

Khan SA, Hussain D, Askari E, Stewart JMD, Malik KA, Zafar Y (2000) Molecular phylogeny of *Gossypium* species by DNA

fingerprinting. Theor Appl Genet 101:931–938. doi:10.1007/s001220051564

Lacape M, Dessauw D, Rajab M, Noyer JL, Hau B (2007) Microsatellite diversity in tetraploid *Gossypium* germplasm : assembling a highly informative genotyping set of cotton SSRs. Mol Breed 19:45–58. doi:10.1007/s11032-006-9042-1

Liu S, Saha S, Stelly D, Burr B, Cantrell RG (2000a) Chromosomal assignment of microsatellite loci in cotton. J Hered 91(4):326–332. doi:10.1093/jhered/91.4.326

Liu S, Cantrell RG, McCarty JCJ, Stewart JMD (2000b) Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. Crop Sci 40:1459–1469

Liu D, Guo X, Lin Z, Nie Y, Zhang X (2006) Genetic diversity of Asian cotton (*Gossypium arboreum* L.) in China evaluated by microsatellite analysis. Genet Resour Crop Evol 53:1145–1152. doi:10.1007/s10722-005-1304-y

Lu HJ, Myers GO (2002) Genetic relationships and discrimination of ten influential Upland cotton varieties using RAPD markers. Theor Appl Genet 105:325–331. doi:10.1007/s00122-002-0947-8

Macaulay M, Ramsay L, Powell W, Waugh R (2001) A representative, highly informative "genotyping set" of barley SSRs. Theor Appl Genet 106:801–809. doi:10.1007/s001220000487

Masi P, Spagnoletti Zeuli PL, Donini P (2003) Development and analysis of multiplex microsatellite markers sets in common bean (Phaseolus vulgaris L.). Mol Breed 11:303–313. doi:10.1023/A:1023443109985

Mehetre SS, Aher AR, Gawande VL, Patil VR, Mokate AS (2003) Induced polyploidy in *Gossypium*: a tool to overcome interspecific incompatibility of cultivated tetraploid and diploid cottons. Curr Sci 84(12):1510–1512

Meredith WR Jr (1991) Contributions of introductions to cotton improvement. In: Shands HL, Weisner LE (eds) Use of plant introductions in cultivar improvement. Part 1 CSSA. Madison, WI, pp 127–146

Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38:1005–1008. doi:10.1139/g95-132

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273. doi:10.1073/pnas.76.10.5269

Park Y-H, Alabady MS, Ulloa M, Sickler B, Wilkins TA, Yu J, Stelly DM, Kohel RJ, El-Shihy OM, Cantrell RG (2005) Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. Mol Genet Genomics 274:428–441. doi:10.1007/s00438-005-0037-0

Reddy OUK, Pepper AE, Abdurakhmonov I, Saha S, Jenkins JN, Brooks TD, Bolek Y, El-Zik KM (2001) New dinucleotide and trinucleotide microsatellite markers for cotton genomic research. J Cotton Sci 5:103–113

Rohlf FJ (2000) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1. User Guide. Exeter Software, New York

Saha S, Wu J, Jenkins JN, McCarty JC, Gutierrez OA Jr, Stelly DM, Percy RG, Raska DA (2004) Effect of chromosome substitutions from *Gossypium barbadense* L. 3–79 into *G.hirsutum* L. TM-1 on agronomic and fiber traits. J Cotton Sci 8:162–169

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Evol Biol 4:406–425

Smulders MJM, Bredemeijer G, Rus-Kortekaas W, Arens P, Vosman B (1997) Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor Appl Genet 97:264–272. doi:10.1007/s001220050409

Song QJ, Quigley CV, Nelson RL, Carter TE, Boerma HR, Strachan JL, Cregan PB (1999) A selected set of trinucleotide simple sequence repeat markers for soybean cultivar identification. Plant Var Seeds 12:207–220

Stewart JMD (1995) Potential for crop improvement through germplasm enhancement and genetic engineering. In: Constable G, Forester N (eds) Challenging the future. Proc World Cotton Research Conference-1. CSIRO, Melbourne, pp 313–327

Struss D, Plieske J (1998) The use of microsatellite markers for detection of genetic diversity in barley populations. Theor Appl Genet 97:308–315. doi:10.1007/s001220050900

Tatineni V, Canterell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci 36:186–192

Ulloa M, Meredith WR Jr, Percy R, Moser H (1999) Genetic variability within improved germplasm of *Gossypium hirsutum* and *G. barbadense* cottons. Crop Sci abstr, Annu Mtg ASA, CSSA, SSSA, Madison, WI, USA, p 73

Ulloa M, Brubaker C, Chee P (2007) Cotton. In: Kole C (ed) Genome mapping & molecular breeding. vol 6: technical crops. Springer, New York, pp 1–49

Van Esbroeck GA, Bowman DT, May OL (1997) Pedigrees and distinguishing features of upland and pima cotton germplasm lines released from 1970 to 1996. N C Agric Res Serv Bull 312

Vergara GV, Bughrara SS (2003) AFLP analyses of genetic diversity in Bentgrass. Crop Sci 43:2162–2171

Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith SC, Doebley J (2002) Rate and pattern of mutation at microsatellite loci in maize. Mol Biol Evol 19:1251–1260

Xu QH, Zhang XL, Nie YC (2001) Genetic diversity evaluation of cultivars (*G. hirsutum* L.) from the Changjiang River valley and Yellow River valley by RAPD markers. Acta Genet Sin 28:683–690

Zhu LF, Zhang XL, Nie YC (2003) Analysis of genetic diversity in upland cotton (*Gossypium hirsutum* L.) cultivars from China and foreign countries by RAPDs and SSRs. J Agric Biotechnol 11:450–455